

# 途上国における数学的リテラシー調査の解答パターンについて PISA2003 の二次分析による先進国との比較から

|     |   |
|-----|---|
| 著者  | 渡邊 耕二   |
| 雑誌名 | 宮崎国際大学教育学部紀要 教育科学論集   |
| 号   | 2   |
| ページ | 14-22   |
| 発行年 | 15-11   |
| URL | <a href="http://id.nii.ac.jp/1106/00000462/">http://id.nii.ac.jp/1106/00000462/</a> |

## 途上国における数学的リテラシー調査の解答パターンについて —PISA2003 の二次分析による先進国との比較から—

渡邊耕二

### The Answer Pattern of Mathematical Literacy Survey in Developing Countries: International Comparison with Developed Countries by Secondary Analysis of PISA2003

Koji Watanabe

#### 1. はじめに

開発途上国（以下、途上国<sup>1</sup>）の数学学力<sup>2</sup>の水準は、国際的にみて低い（e.g., OECD、2004、2007、2010 ; IEA、2005、2008a、2008b）。このことは、途上国における教育の質向上の必要性を唱える根拠の一つとなっている。

一般に、途上国の低い学力水準を表わす指標として、複数の出題項目（以下、項目）の得点を総合して得られるテスト得点が用いられる。しかし、テスト得点に注目したときに、同程度の学力水準であったとしても、各項目の正答率などが各国で同様であるとは限らない。そのためテスト得点だけでなく、各項目の特性に目を向け、解答パターンという視点から途上国の特徴を捉える必要がある。

解答パターンを捉える際には、各項目の正答率がよく用いられる。例えば、国立教育政策研究所（2004、2013）では、日本を含む13カ国におけるいくつかの項目の正答率が提示されている。しかしながら正答率には、ある特定の受検者集団に依存して定められるという性質がある（芝、1991）ため、各項目の難易度といった特性を客観的に比較しにくい。解答パターンをより詳細に捉えるためには、受検者集団と出題項目の特性を区別する必要がある。このような分析を行う方法として、現在では、項目反応理論が知られている（e.g., 芝、1991 ; 豊田、2002 ; 加藤・山田・川端、2014）。

鈴木・豊田・川端（2008）は、国立教育政策研究（2004）で取り上げられた13カ国を対象に、項目反応理論で定義される困難度と呼ばれる項目の難易度に着目し、PISA2003のデータを二次分析している。そして日本は、日常場面に近い状況<sup>3</sup>の項目において独特な解答パターンを持つことを示している。また渡邊（2012a）は、同じ13カ国を対象にPISA2003を用いて、日本の解答パターンを確率・統計<sup>4</sup>の領域を中心に分析し、日本の生徒は、グラフにある数値の読み取りに難点を有すると指摘した。

このように、日本をはじめいくつかの国において、PISA2003の数学的リテラシー調査のデータを用いて、解答パターンが調べられている。しかしながら、低い学力水準の国に

<sup>1</sup> 本研究では、OECD 開発援助委員会（Development Assistance Committee : 通称、DAC）援助受取国・地域リスト 2007 年度版（外務省、2008）に挙げられた国々を途上国と規定する。

<sup>2</sup> 「学力」という言葉は、定義が難解な用語である（荻谷・志水、2004 ; 市川、2001）といわれるが、本稿では、テストによって浮かび上がる測定可能な力として「学力」を規定する。

<sup>3</sup> PISA2003 で規定された「数学が用いられる状況」に着目し、考察している。

<sup>4</sup> PISA2003 で規定された「包括的アイディア」に含まれる「不確実性」に着目し、考察している。

については、必ずしも目が向けられていない。そこで本研究では、低い学力水準と称される途上国に目を向けて PISA2003 のデータを用いて解答パターンを検討する。

## 2. 研究方法

### 2.1. 使用データとデータ処理

本研究では、PISA2003 の公開データを使用する。分析対象は、実施基準を満たしていないと判断されたイギリスと受検者数が 332 名と特に少ないリヒテンシュタインを除く 39 カ国の 265250 名である。なお PISA2003 では、数学的リテラシー<sup>5</sup>を測定するために、13 種類のブックレットが使用されている<sup>6</sup>。しかしそれとは別に、短い時間で解答できる UH ブックレットに 1048 名が解答しているが、人数が少ないためこれらの受検者は分析対象に含めていない。

数学的リテラシーの測定に使用された各項目のデータ処理については、OECD (2005) のコードブックをみながら、渡邊 (2012a) と同じ処理を行った<sup>7</sup>。具体的には、無回答、Not reached、部分点、欠損値は誤答として扱うことにした。図 2.1 に実際に使用したデータセットの 10 名分を提示した。NA は、受検者が解答したブックレットに含まれていない項目である。

図 2.1 データ処理後のデータセットの一例

| 国名  | ブックレット | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 | ... |
|-----|--------|----|----|----|----|----|----|----|----|----|-----|-----|
| AUS | 11     | 0  | NA | NA | NA | NA | NA | NA | NA | NA | NA  | ... |
| AUS | 9      | NA | NA | NA | NA | NA | NA | NA | NA | NA | 1   | ... |
| AUS | 12     | NA | 1  | NA | NA | NA | NA | NA | NA | 1  | NA  | ... |
| AUS | 1      | 1  | 1  | NA | NA | 1  | 0  | 1  | 0  | 1  | NA  | ... |
| AUS | 9      | NA | NA | NA | NA | NA | NA | NA | NA | NA | 0   | ... |
| AUS | 7      | NA | NA | 0  | 0  | NA | NA | NA | NA | NA | NA  | ... |
| AUS | 5      | 1  | NA | NA | NA | NA | NA | NA | NA | NA | 1   | ... |
| AUS | 13     | 1  | NA | 0  | 0  | NA | NA | NA | NA | NA | NA  | ... |
| AUS | 6      | NA | 1  | NA | NA | NA | NA | NA | NA | 1  | NA  | ... |
| AUS | 2      | NA | 0  | 0  | 0  | NA | NA | NA | NA | 1  | 1   | ... |

### 2.2. 分析の流れ

まず、受検者のテスト得点を算出し、各国の数学的リテラシーの水準を確認する。テスト得点の計算には、項目反応理論における項目反応モデルの一つであるラッシュモデル<sup>8</sup>を採用する。

次に、算出したテスト得点の平均値を降順に並べ、39 カ国を 4 つのグループに分ける。そのグループごとに解答パターンの特徴を調べ、学力水準が高い国と低い国における解答

<sup>5</sup> PISA の「数学的リテラシー」は、「数学が現実で果たす役割りを見つけ、理解し、現在および将来の個人の生活、職業生活、友人や家族や親族との社会生活、建設的で関心を持った思慮深い市民としての生活において確実な根拠に基づき判断を行い、数学に携わる能力」と規定される（国立教育政策研究所、2004）。

<sup>6</sup> PISA では、重複テスト分冊法が採用されており、各項目には、全体の約 3 割の受検者が解答するように設定されている。

<sup>7</sup> 渡邊 (2012b、2014、2015) のデータ処理とも共通している。

<sup>8</sup> ラッシュモデルは、PISA でも採用されている。このモデルは、項目母数が困難度の一つであるため、扱いが容易である（豊田、2012、p.25）。

パターンの違いを検討する。なお、各グループにおける解答パターンの特徴を調べる方法は、鈴川・豊田・川端（2008）を参考にし、ラッシュモデルに含まれる項目の難易度に関するパラメータ（以下、困難度）に着目する。

最後に、渡邊（2012b）が取り上げた 7 途上国、具体手にはトルコ、タイ、ウルグアイ、メキシコ、ブラジル、インドネシア、チュニジアに目を向け、途上国における解答パターンの特徴を検討する。

### 3. 分析結果

#### 3.1. 数学的リテラシー調査のテスト得点の国際比較

ラッシュモデルを用いてテスト得点を算出する場合には、各項目の困難度の絶対値が 4.0 以下になることが望ましいとされる（芝、1991：豊田、2002）。そこで、テスト得点を計算するために、表 3.1 に使用する項目の困難度の基本統計量を示した。各項目の絶対値は、4.0 以下であるため、84 項目を用いてテスト得点を算出する。

表3.1 困難度の  
基本統計量

|      |        |
|------|--------|
| 項目数  | 84     |
| 平均値  | 0.275  |
| 標準偏差 | 0.778  |
| 最大値  | 2.214  |
| 最小値  | -1.842 |
| レンジ  | 4.056  |

表 3.2 に各国のテスト得点の分布を平均値と標準偏差を用いて提示した。なお受検者全体で平均値 0.0、標準偏差 1.0 と調節してある。表 3.2 をみると上位層には、日本を含む先進国が位置し、途上国は下位層に位置することが分かる。

#### 3.2. 数学学力の水準の違いに着目した解答パターンの国際比較

テスト得点から各国の数学学力の水準を確認した。しかし学力水準が同程度であったとしても、国内における解答パターンが同様であるとは限らない。そこで表 3.2 に示したように、テスト得点の平均値の降順に 39 カ国を 4 つのグループに分け、グループごとに解答パターンの特徴を困難度に着目して調べる。なお PISA2003 の数学的リテラシー調査では、アイスランド、韓国、イタリア、デンマーク、ハンガリー、カナダ、スイス、ウルグアイの 8 カ国において、整合性が低いと判断された項目が示されている<sup>9</sup>。8 カ国を含むグループでは、該当する項目を分析対象外とし、各グループで項目数を揃えて分析を行う。なお各グループにおける分析対象とした項目数は、グループ 1 では 83 項目、グループ 2 では 80 項目、グループ 3 では 82 項目、グループ 4 では 82 項目である。

項目反応理論を用いると、複数の受検者集団から得られる困難度を等化し、比較可能な数値へと変換できる。本分析では、mean-sigma 法<sup>10</sup>を用いて、グループ 1 は香港、グループ 2 はスイス、グループ 3 はノルウェー、グループ 4 はギリシアにそれぞれ等化を施す。この方法によると、各国の困難度の平均値がグループごとに同一の値に揃えられる<sup>11</sup>。さらに、 $b_{ij}$  を  $j$  ( $j=1,2,\dots,m$ ) 国の項目  $i$  ( $i=1,2,\dots,n$ ) の等化後の困難度とし、各国

<sup>9</sup> OECD（2005、p.190）によると、アイスランド（M144Q03）、韓国（M155Q01）、イタリア（M179Q01T）、デンマーク（M273Q01）、ハンガリー（M402Q02）、カナダ（M603Q02）、スイス（M704Q01T）、ウルグアイ（M442Q02 と M800Q01）に整合性の低い項目が指摘されている。

<sup>10</sup> 豊田（2012）が詳しい。

<sup>11</sup> 具体的には、各グループの基準国の平均値に揃えられる。

表 3.2 数学的リテラシーのテスト得点の分布（平均値の降順）

| 国        | 平均値    | 標準誤差  | 標準偏差  | 標準誤差  | 受検者数   |
|----------|--------|-------|-------|-------|--------|
| グループ1    |        |       |       |       |        |
| 香港       | 0.627  | 0.014 | 0.942 | 0.019 | 4478   |
| オランダ     | 0.550  | 0.014 | 0.877 | 0.017 | 3908   |
| フィンランド   | 0.527  | 0.011 | 0.854 | 0.014 | 5796   |
| 韓国       | 0.507  | 0.012 | 0.900 | 0.016 | 5444   |
| ベルギー     | 0.492  | 0.011 | 0.986 | 0.015 | 8549   |
| チェコ      | 0.479  | 0.012 | 0.949 | 0.016 | 6192   |
| 日本       | 0.453  | 0.014 | 0.968 | 0.019 | 4707   |
| ニュージーランド | 0.375  | 0.014 | 0.959 | 0.019 | 4511   |
| オーストラリア  | 0.351  | 0.008 | 0.952 | 0.011 | 12551  |
| マカオ      | 0.338  | 0.026 | 0.910 | 0.033 | 1250   |
| グループ2    |        |       |       |       |        |
| スイス      | 0.332  | 0.010 | 0.937 | 0.014 | 8420   |
| カナダ      | 0.331  | 0.005 | 0.890 | 0.007 | 27953  |
| フランス     | 0.282  | 0.014 | 0.911 | 0.018 | 4300   |
| アイスランド   | 0.280  | 0.016 | 0.915 | 0.020 | 3350   |
| デンマーク    | 0.278  | 0.014 | 0.920 | 0.018 | 4218   |
| オーストリア   | 0.264  | 0.013 | 0.885 | 0.016 | 4568   |
| ドイツ      | 0.260  | 0.014 | 0.934 | 0.018 | 4552   |
| スウェーデン   | 0.226  | 0.014 | 0.939 | 0.018 | 4624   |
| スロバキア    | 0.196  | 0.011 | 0.909 | 0.014 | 7288   |
| アイルランド   | 0.187  | 0.014 | 0.854 | 0.017 | 3880   |
| グループ3    |        |       |       |       |        |
| ノルウェー    | 0.119  | 0.015 | 0.929 | 0.019 | 4064   |
| イタリア     | 0.116  | 0.009 | 0.926 | 0.011 | 11639  |
| ハンガリー    | 0.108  | 0.014 | 0.907 | 0.018 | 4371   |
| ルクセンブルク  | 0.101  | 0.014 | 0.899 | 0.018 | 3923   |
| スペイン     | 0.101  | 0.008 | 0.875 | 0.010 | 10791  |
| ポーランド    | 0.052  | 0.014 | 0.910 | 0.018 | 4383   |
| ラトビア     | 0.021  | 0.013 | 0.880 | 0.016 | 4627   |
| アメリカ     | -0.028 | 0.012 | 0.900 | 0.015 | 5456   |
| ロシア      | -0.088 | 0.012 | 0.932 | 0.016 | 5974   |
| ポルトガル    | -0.153 | 0.013 | 0.872 | 0.016 | 4608   |
| グループ4    |        |       |       |       |        |
| ギリシア     | -0.343 | 0.013 | 0.909 | 0.017 | 4627   |
| 旧ユーゴスラビア | -0.382 | 0.013 | 0.839 | 0.015 | 4405   |
| トルコ      | -0.466 | 0.013 | 0.930 | 0.018 | 4855   |
| タイ       | -0.510 | 0.012 | 0.843 | 0.014 | 5236   |
| ウルグアイ    | -0.542 | 0.012 | 0.945 | 0.017 | 5835   |
| メキシコ     | -0.647 | 0.005 | 0.779 | 0.005 | 29983  |
| ブラジル     | -0.958 | 0.012 | 0.833 | 0.015 | 4452   |
| インドネシア   | -0.982 | 0.007 | 0.719 | 0.007 | 10761  |
| チュニジア    | -0.993 | 0.011 | 0.746 | 0.011 | 4721   |
| 全体       | 0.000  | 0.002 | 1.000 | 0.003 | 265250 |

の項目  $i$  の困難度を用いて、グループごとに項目  $i$  の平均値  $\bar{b}_i$  を求める。 $b_{ij}$  と  $\bar{b}_i$  の偏差  $b_{ij}^* = b_{ij} - \bar{b}_i$  を考えると、各国の困難度の平均値  $\bar{b}_j^*$  および各グループの項目  $i$  の平均値  $\bar{b}_i^*$  は、それぞれ 0.0 となる。つまりグループごとに、各国と各項目の平均値を 0.0 に統一し、各グループの特徴を浮かび上がらせる<sup>12</sup>。なお分析対象とする困難度は、グループ 1 では 83 項目×10 カ国、グループ 2 では 80 項目×10 カ国、グループ 3 では 82 項目×10 カ国、

<sup>12</sup> 鈴川・豊田・川端（2008）が詳しい。

グループ 4 では 82 項目×9 カ国の合計 3188 になるため、結果の提示は割愛する。

数学的リテラシー調査の項目は、①包括的アイディア（数学的な内容）、②数学のプロセス（能力クラスター）、③数学が用いられる状況の 3 つの側面から特徴付けられている。特に、数学が用いられる状況（以下、状況・文脈）は、数学的リテラシーを真正に評価するために設定されている<sup>13</sup>。本分析では、状況・文脈に着目し、各グループの特徴を調べる。なお、表 3.3 に状況・文脈にある分類を提示した。

各グループの特徴を調べるために、グループごとに主成分分析を適用し、第 1 主成分と第 2 主成分を軸とする各項目の主成分得点の散布図を作成した。図 3.1、図 3.2、図 3.3 および図 3.4 は、それぞれグループ 1、グループ 2、グループ 3 およびグループ 4 における散布図である。各図にある「私」は私的、「教」は教育的、「職」は職業的、「公」は公共的、「科」は科学的に分類された項目を表わす。また、各グループにおける第 2 主成分までの累積寄与率は、グループ 1 では 0.556、グループ 2 では 0.508、グループ 3 では 0.431、グループ 4 では 0.415 である。図 3.4 から分かるように、グループ 4 において Intra-M は、明らかに突出した傾向を示している。つまり、数学学力の水準が低いグループ 4 においては、数学そのものの文脈に関する項目が目立った特徴が浮かび上がった。

表 3.3 状況・文脈の分類と内容<sup>14</sup>

| 分類      | 内容  |
|---------|---|
| 私的      | 生徒の日々の活動に直接関係する文脈                               |
| 教育的     | 生徒の学校生活に現れるような文脈                                |
| 職業的     | 職業の場面に現れるような文脈                                  |
| 公共的     | 生徒が生活する地域社会における文脈                               |
| 科学的     | より抽象的な文脈で、技術的な過程、理論的な場面、明らかに数学的な問題についての理解に関する文脈 |
| Intra-M | 数学の授業でよく直面するような数学そのものの文脈                        |

出典：国立教育政策研究所（2004、p.34）を参考に筆者作成

### 3.3. 7 途上国に着目した分析

39 カ国を数学学力の水準別に 4 つのグループに分け、解答パターンの特徴を調べた。その結果、低い学力水準のグループでは、数学そのものの文脈に関する項目が特異であることが分かった。そこで、渡邊（2012）が取り上げた 7 カ国に目を向け、解答パターンの特徴をみていく。この 7 カ国に対して、4 つのグループに施した同様な分析と行ったところ、図 3.5 の結果を得た。なお、第 2 主成分までの寄与率は 0.459 である。図 3.5 をみると、図 3.4 と同様に Intra-M に目立った特徴を確認できる。またそれは、図 3.4 よりも原点から離れた位置にあると読み取れるため、より突出した特徴を持つと考えられる。

<sup>13</sup> 学校の教科書に煩雑にみられる数学を練習することではなく、様々な状況において数学を用いて問題を解決できるかをみるために設定されている（国立教育政策研究所、2004、p.33）。

<sup>14</sup> Intra-M は、「数学内的」として科学的に含まれる項目の一つであるが、本稿では区別して扱っている。

途上国における数学的リテラシー調査の解答パターンについて  
—PISA2003 の二次分析による先進国との比較から—

図 3.1 グループ 1 の主成分得点の散布図

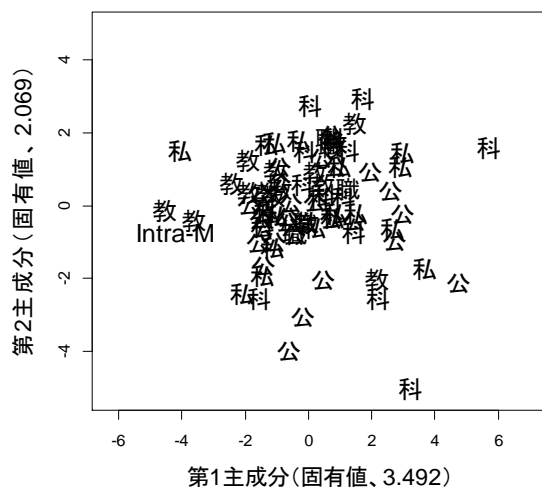


図 3.2 グループ 2 の主成分得点の散布図

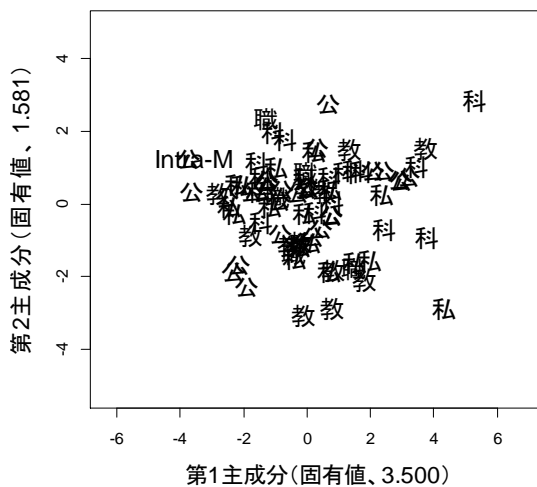


図 3.3 グループ 3 の主成分得点の散布図

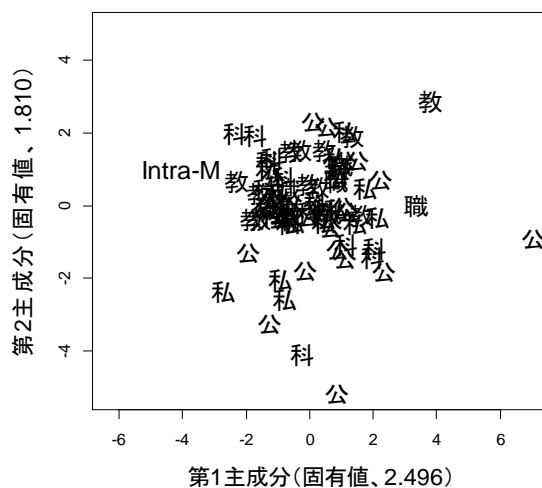


図 3.4 グループ 4 の主成分得点の散布図

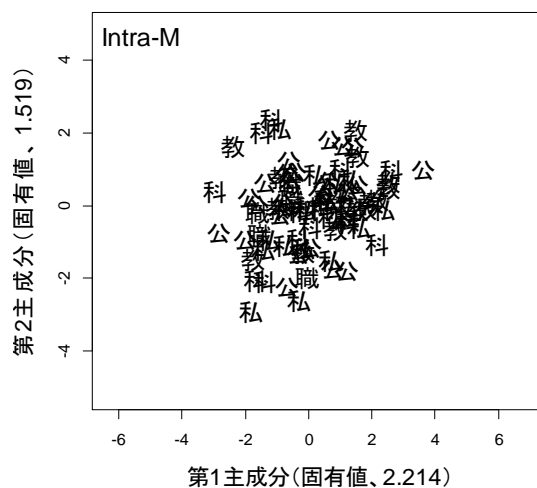
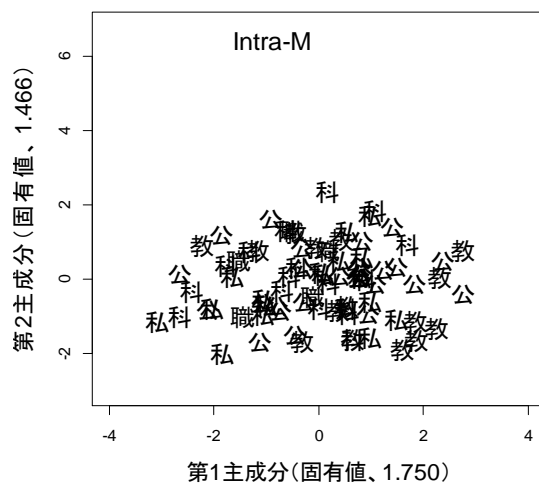


図 3.5 7 カ国における主成分得点の散布図



#### 4. おわりに

PISA2003 の公開データを二次分析し、国際的にみて数学学力の水準が低い国に目を向け、数学的リテラシー調査における解答パターンを調べた。

一般に、途上国の数学学力の水準は低い。本分析においても、その傾向を確かめた（表 3.2）。しかし、学力水準が同程度であったとしても、国内における解答パターンが同様であるとは限らない。そこで、39 カ国を学力の水準別に 4 つのグループに分け、グループごとの解答パターンの特徴を分析した。その結果、他のグループとは異なり、グループ 4 における Intra-M は、他の項目と比べて特異なことが分かった（図 3.1、図 3.2、図 3.3、図 3.4）。また 7 途上国に絞って分析したところ、その特徴がより顕著に現れると考えられる（図 3.5）。纏めると Intra-M は、日本を含む学力水準が高い国では特異性を示さないが、逆に学力水準が低い途上国においては、突出した特徴を有するといえるだろう。

近年では、途上国の低い数学学力の向上に向けて、数学的な知識を身に付けることや計算を正しく行うことだけでなく、知識や技能を応用する力の育成が求められている（国際協力機構、2007、p.29）。しかしながら本分析から途上国では、授業でよく直面する数学に対する内容において特異な解答パターンを持つことが示された。つまり、途上国を低い学力水準だけでなく、解答パターンの視座から特徴付けるならば、知識や技能を応用するための数学的な知識や計算という側面に十分に目を向けるべきと考える。

これまで、途上国に関する数学教育研究では、PISA や TIMSS といった大規模な学力調査のデータがほとんど活用されてこなかった（河原、2014）。また、途上国の低い学力水準のみが強調され、有益な情報を得ることが難しいとされてきた（馬場・内田、2008；内田、2009）。しかし本稿では、低い学力水準だけでなく、項目反応理論の困難度に着目し、途上国の特徴を解答パターンの視座から一定程度浮かび上がらせた。その意味で本分析の成果は、途上国の数学教育研究に一石を投じるものである。

なお本研究は、PISA2003 の公開データを分析したが、PISA2006、PISA2009 および PISA2012 とデータの蓄積が進んでいる。蓄積されたデータを分析対象とし、経年変化を含めて途上国の特徴を捉える取り組みが今後必要である。また、国際的な学力調査として PISA だけでなく TIMSS も知られており、同様にデータが蓄積され、ウェブ上に公開されている。TIMSS は、PISA とは異なる尺度で数学学力を測定しており、これらのデータを活用し、多面的に途上国の特徴を捉える必要もあるだろう。これらのデータを用いた分析は、今後の課題としたい。

#### 【謝辞】

本研究は、JSPS 科研費 26885120 の助成を受けて行われたものである。

#### 参考文献

- IEA. (2005). *IEA's TIMSS2003 International Report on Achievement in the Mathematics Cognitive Domains*, IEA and Boston College.
- IEA. (2008a). *TIMSS2007 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*, IEA and Boston College.



- IEA. (2008b). *TIMSS2007 International Science Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*, IEA and Boston College.
- OECD. (2004). *Learning for Tomorrow's World: First Results from PISA2003*, OECD: Paris.
- OECD. (2005). *PISA2003 Technical Report*, OECD: Paris.
- OECD. (2007). *Learning for Tomorrow's World: First Results from PISA2006*, OECD: Paris.
- OECD. (2010). *PISA2009 Results: What Students Know and Can Do –Student Performance in Reading, Mathematics and Science (Volume 1)*, OECD: Paris.
- 馬場卓也・内田豊海. (2008). 「国際比較調査の開発途上国の教育開発に対するインプリケーションに関する考察—ガーナ国 TIMSS レポートを事例として—」, 『国際教育協力論集』, 第 11 号 第 2 号, pp.129-140.
- 外務省. (2008). 『2007 年版政府開発援助 (ODA) 白書—日本の国際協力—』, 外務省.
- 市川伸一. (2001). 「学力低下論争の構図と『もう一つの学力低下論』」, 中井浩一編『論争・学力崩壊』, 中公新書ラクレ, pp.209-231.
- 荻谷剛彦・志水宏吉. (2004). 『学力の社会学—調査が示す学力の変化と学習の課題』, 岩波書店.
- 加藤・山田・川端. (2014). 『R による項目反応理論』, オーム社.
- 河原太郎. (2014). 「TIMSS・PISA の 2 次分析に関する英語論文の実態—Eric の Journal Article に焦点を当てて—」, 『国際協力研究誌』, 第 20 巻 第 2 号, pp.1-7.
- 国立教育政策研究所. (2004). 『生きるための知識と技能 2—OECD 生徒の学習到達度調査 PISA2003 年調査国際結果報告書』, ぎょうせい.
- 国立教育政策研究所. (2013). 『生きるための知識と技能 5—OECD 生徒の学習到達度調査 PISA2012 年調査国際結果報告書』, 朝倉書店.
- 国際協力機構. (2007). 『理数科教育協力にかかる事業経験体系化—その理念とアプローチ—』, 国際協力機構.
- 芝祐順. (1991). 『項目反応理論 基礎と応用』, 東京大学出版会.
- 鈴川由美・豊田秀樹・川端一光. (2008). 「わが国の数学教育は数学を日常の中で活用する能力を重視しているか—PISA2003 年調査の DIF による分析—」, 『教育心理学研究』, 第 56 巻 第 2 号, pp.206-217.
- 豊田秀樹. (2002). 『項目反応理論 [入門編]—テストと測定の科学—』, 朝倉書店.
- 豊田秀樹. (2012). 『項目反応理論 [入門編] (第 2 版)』, 朝倉書店.
- 内田豊海. (2009). 「ザンビア基礎教育における計算能力の診断的評価に関する研究—弁別性と教授的示唆に注目して—」, 『国際教育協力論集』, 第 12 巻 第 2 号, pp.1-12.
- 渡邊耕二. (2012a). 「わが国の生徒が持つ確率・統計を活用する能力の特徴に関する研究—PISA2003 「数学的リテラシー」 の項目反応理論による二次分析から—」, 『数学教育学研究』, 第 18 巻 第 1 号, pp.59-70.
- 渡邊耕二. (2012b). 「わが国の生徒が持つ数学に対する情意的側面と認知的側面の関連性

について—PISA2003 の質問紙調査と数学的リテラシー調査の二次分析から—,『日本数学教育学会誌』, 第 94 巻 第 11 号, pp.12-21.

渡邊耕二. (2014). 「開発途上国における数学の認知的側面と情意的側面の関連性に関する研究—PISA2003 の二次分析を通じて—」,『数学教育学研究』, 第 20 巻 第 2 号, pp.63-76.

渡邊耕二. (2015). 「数学学力と読解力に着目した言語的な側面の関係性に関する国際比較—PISA2003 と PISA2012 の二次分析から—」,『数学教育学研究』, 第 21 巻 第 2 号, pp.73-87.